Televisions, Video Privacy, and Powerline Electromagnetic Interference

Miro Enev
University of Washington
miro@cs.washington.edu

Sidhant Gupta
University of Washington
sidhant@cs.washington.edu

Tadayoshi Kohno
University of Washington
yoshi@cs.washington.edu

Shwetak N. Patel
University of Washington
shwetak@cs.washington.edu

ABSTRACT

We conduct an extensive study of information leakage over the powerline infrastructure from eight televisions (TVs) spanning multiple makes, models, and underlying technologies. In addition to being of scientific interest, our findings contribute to the overall debate of whether or not measurements of residential powerlines reveal significant information about the activities within a home. We find that the power supplies of modern TVs produce discernible electromagnetic interference (EMI) signatures that are indicative of the video content being displayed. We measure the stability of these signatures over time and across multiple instances of the same TV model, as well as the robustness of these signatures in the presence of other noisy electronic devices connected to the same powerline.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection—*Unauthorized access*; I.5.4 [**Pattern Recognition**]: Applications—*Signal Processing*; I.5.1 [**Pattern Recognition**]: Models—*Neural Nets*

General Terms

Security, Experimentation

Keywords

Information leakage, powerline security, electromagnetic interference

1. INTRODUCTION

It is known that consumer electronic devices can leak *some* information over powerlines, such as whether a particular device is on or off [10, 23], or even information about specific keystrokes on a keyboard [3, 28]. However, this knowledge is limited. In many cases, when faced with a given technology and asked whether it leaks private information over the powerline, our only possible answer is either a "maybe" or a "probably." The device probably leaks information about whether it is on or off. It might also leak more fine-grained information, but exactly what information we do not know. Moreover, even if we can make an educated guess that a device might leak some particular form of information, we do not know how stable that information leakage channel is over time or over multiple instances of the same technology, nor how robust that information leakage is against interference from other devices connected to the same powerline.

In tandem with the above observations, we find that numerous technologies are being developed that will measure information from residential powerlines. For example, utility (power) companies are seeking to deploy smart meters that measure fine-grained power usage for billing and demand-prediction [27], and ubiquitous computing (ubicomp) systems similarly measure the powerline for activity recognition purposes [10, 11, 23]. Future versions of these technologies may measure even more information. Returning to the above-mentioned information leakage knowledge gap, there is enough information today for privacy advocates to argue that these powerline measurement should be treated as private [20], thus motivating proposed policy changes, recommendations to industry, and research on privacy-preserving mechanisms for powerline measurements (e.g., [24]). Those opposed to these recommendations argue that the privacy concerns are overinflated. Unfortunately, the existence of the knowledge gap means that both sides of the debate are proceeding with incomplete knowledge and hence positions based in part on conjecture. Thus, in addition to being of scientific interest, new research is needed to inform this debate.

While it is infeasible to completely close this knowledge gap and study *all* electronic devices and root out *all* powerline information leakage channels, it is feasible — even if daunting — to rigorously and extensively study individual artifacts or classes of artifacts. For this study, we choose to focus extensively on one of the most ubiquitous classes of consumer electronic devices: modern televisions (TVs). We choose TVs, rather than toasters or ovens, because TVs have complex internal states — functions of the screen content being displayed — and thus present the *potential* for rich information leakage.

Separately, our choice of TVs was driven by the legal status of TVs and video viewing habits. Specifically, the U.S. Video Protection Privacy Act of 1998 states that an

CCS'11, October 17–21, 2011, Chicago, Illinois, USA.

Television Name	Technology	Date of Fabrication	Price (USD)
Panasonic-42-A	Plasma	Sept 2010	$1099
Panasonic-42-B	Plasma	Sept 2010	$1099
Samsung-58-A	Plasma	Sept 2010	$2599
Samsung-58-B	Plasma	Sept 2010	$2599
Samsung-32	LCD	Oct 2007	$896
Sharp-42	LCD	Sept 2008	$1399
Sharp-32-A	LCD	Dec 2009	$499
Sharp-32-B	LCD	Dec 2009	$499

Table 1: A summary of the tested TVs showing make, technology, date of manufacture, and price for each device.

individual's video viewing records must be kept confidential. Any research finding to the contrary — that a measurements of a home's powerline could reveal private information about someone's video viewing habits — could thus have significant legal ramifications.

For breadth and depth, we obtained a collection of eight TVs spanning three manufacturers (Panasonic, Samsung, and Sharp),[1] two technologies (LCD and Plasma), and three sizes (32", 42", and 58"); see Table 1. We repeatedly played video sequences on each of these TVs while recording the electromagnetic interference (EMI) that these TVs produced as measured on the powerlines; we obtained these measurements using a separate single sensor attached to the same powerline, though explicitly *not* between the TV and the powerline itself. Our sensor's access to the powerline thus matches the access afforded to any other consumer electronic plugged in somewhere within a home. We analyzed the EMI in the frequency domain. We obtained initial measurements in a clean laboratory setting, and then collected further measurements in unrestricted home environments with diverse collections of other noisy electronics attached to the same powerline.

We find that all but the Sharp 32″ TVs produce a significant amount of stable, robust information leakage over the powerline via EMI. As an example, for the Panasonic 42″ TVs in a laboratory setting, we find that we can correctly match a random 15 minute EMI trace from a movie to a database of EMI signatures totaling 1200 movie minutes 96% of the time (the remaining 4% were below our matching threshold, not incorrectly matched); the hit rate drops to 92% in noisy home environments. More broadly, we sought to develop an infrastructure capable of exploring answers to the following five questions:

- Do TVs of different makes, models, and technologies produce repeatable EMI, when given repeated screen content?
- Does different screen content produce differentiable EMI from a given TV?
- Is EMI consistent across multiple instances of TVs from the same model family?
- Can we determine if someone is watching a particular video by matching an EMI sample to a database of EMI signatures?
- Can we match EMI recorded in a lab setting to EMI extracted in homes, especially when those homes might have a diverse array of other consumer electronics connected to the same powerline?

In general, we find that we are able to answer all these questions affirmatively, though with some caveats and clarifications, e.g., clarifications on what it means for an EMI trace to be "repeatable" or "consistent" and weaker results with the Sharp 32″ TVs. Although explicitly not our primary goal, we also sought to conduct a preliminary investigation into the following question:

- Can we learn to predict what EMI a TV would produce given *only* screen content, i.e., without previous EMI traces for that TV model when displaying the content in question?

To summarize, our primary contribution is an extensive study of powerline EMI information leakage from a diverse collection of modern TVs, as well as the discovery that significant, stable, and robust information is, in fact, leaked from these devices. Fortunately, from a privacy perspective, we also observe that today's utility companies are *not* collecting the granularity of information necessary for repeating our analyses, and the ubicomp technologies from the research community that could collect this type of information are not yet widely deployed. Nevertheless, these technologies are surprisingly near-term, e.g., smart meters exist today that report samples at 1 Hz but collect measurements at 3 kHz [7] (our approach in contrast samples at 500 kHz and uses measurements at 2 Hz) and we anticipate higher sampling rates in the near future.

Although these emerging technologies enable important applications, we stress here that the policies enacted today could help inform future utility and ubicomp deployments, and hence we argue that it is important to understand these findings today, before new technologies in or attached to the home begin to collect and store fine-grained powerline measurements at a large scale.

2. TECHNICAL FOUNDATIONS

Modern consumer electronics utilize switched mode power supplies (SMPS), which are increasingly employed due to their higher efficiency and smaller form factor. Energy Star even mandates the use of this kind of power supply for most consumer electronic devices. We anticipate that the prevalence of SMPS will only increase as hardware manufacturers face continued pressure from political and market forces demanding higher levels of efficiency and miniaturization. The materials described here are either known or easily inferred from past work and an understanding of how electrical devices operate, e.g., [10].

SMPS as EMI Sources. Despite the advantages SMPS provide, one side effect of SMPS's operation is that they inadvertently produce electromagnetic interference (EMI). The EMI is a consequence of the voltage conversion method that SMPS employ and it's characteristics are proportional to the rate at which the conversion in the power supply takes place. This rate, also known as the switching frequency, is dynamically adjusted to meet the power requirements of the appliance driven by the SMPS. In sum, the changing power demands of an electronic device cause dynamic adjustments to the SMPS's switching frequency which in turn modulates

[1]We do not have any reason to believe that our results are specific to these manufacturers.

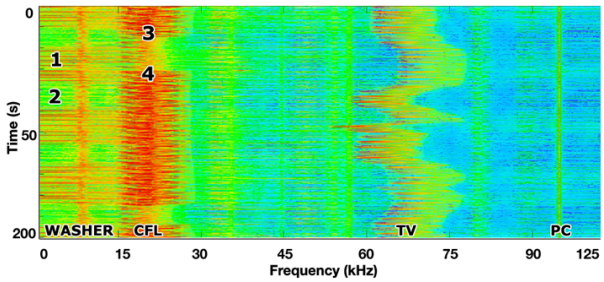


Figure 1: Frequency spectrogram showing various electrical appliances in the home. Washer cycle on (1) and off (2). CFL lamp turning off briefly (3) and then on (4). Note that the TV's (Sharp 42" LCD) EMI shifts in frequency, which happens as screen content changes.

the EMI signal. To prevent high levels of inadvertent EMI from adversely influencing the operation of neighboring devices, the Federal Communications Commission (FCC) in the U.S. limits the amount of interference that a device can emit to -40dBm for the frequency range of 1–500 kHz (47CFR part 15/18 Consumer Emission Limits); however the regulation threshold is weak from an information leakage perspective since EMI measurements can be accurately done using inexpensive hardware (the circuit we describe is sensitive to -100 dBm).

Noise Propagation. The EMI generated by SMPS devices is coupled onto the powerline and propagates throughout the electric infrastructure (i.e., a home) while retaining its signal qualities. In our tests we were able to capture robust signals several hundred feet away from the EMI source. We leverage this property and demonstrate that a single sensor anywhere along a home's powerline (including outdoor outlets) can observe the conducted EMI signals from numerous connected SMPS-based devices.

Signal Characteristics. The raw signals we gather with our powerline interface (PLI) (detailed in Section 5) are not particularly meaningful unless analyzed in the frequency domain. To transform the signal from a sequence of data points in time to a set of energies at various frequencies, we digitize the analog readings and compute a Fast Fourier Transform (described in detail later in the paper). The goal of these engineering operations is to allow us to analyze EMI fluctuations as changes in power over frequency space (visualized in Figure 1).

Figure 1 presents a waterfall plot of the EMI observed in a typical home; note the rich information landscape which unfolds when the frequencies of EMI signals (horizontal axis) are plotted against time (vertical axis). In this snapshot we observe that several devices are in continuous operation (PC, television, and washer) while others are turning on and off (compact fluorescent lamp (CFL)).

There are several interesting features in this data, which provide a good reference point for discussing the general EMI signal characteristics that we observed across our many experiments. In particular:

- Each SMPS device produces some conducted EMI, and the presence or absence of this signal is a direct consequence of the power state of the device.
- The EMI signature of each device is centered around the switching frequency of its power supply; different devices often occupy distinct frequency ranges allowing for their noise signals to coexist with minimal interference. However signal overlap is possible, and increasingly likely if many active devices share the same powerline.
- Changes in EMI can manifest in frequency shifts, energy density changes, or both. These properties are visible in Figure 1 as the TV exhibits frequency shift EMI, while the washer signal produces density changes.

Note that EMI signals vary as a function of the operation of the device, e.g., washer EMI increases in energy density when the barrel is rotated during the spin cycle, and TV EMI is modulated in frequency as the screen content changes. We explore this phenomenon more deeply in our investigations below.

3. RELATED WORK

Our work builds on and complements past work in the activity recognition, energy monitoring, and information leakage.

Activity Recognition & Energy Monitoring. Detecting electrical device activity and power consumption in the home has historically been done in the distributed sensing model wherein each device being tracked is instrumented with a dedicated sensor. This one-sensor-per-device model is limiting because, as the name suggests, every monitored device requires separate sensor installation and maintenance. However, research dating from the 1980s in single-point sensing of electrical events has focused on the design of new in-line metering techniques for monitoring whole-house appliance usage [11]. This early work used a current sensor installed in-line with a home's power meter to monitor voltage and current waveforms from the incoming power to the home. The methods used in these original studies combined measurements of (1) step changes in power, (2) the active and reactive power of the home, and (3) time of day statistics. These data sources were combined to detect and classify individual appliance usage.

Recent advances in infrastructure-mediated sensing (IMS) from the ubicomp literature have provided an alternative approach for detection of events on the home's utility infrastructure [6, 9, 10, 22, 23]. These methods focus on leveraging incidental signals such as the EMI from electrical appliances and water pressure changes from plumbing fixtures for inferring appliance usage and thus in-home activity from a single-point. We were inspired by this line of work and have based our sensor architecture on the design of a power line interface introduced in [10] to capture EMI present on a home's powerline. The primary focus of the existing research on electrical sensing has been on detecting and differentiating between appliances, whether by measurements of overall power consumption [11] or via the presence/absence of characteristic signals in the EMI when analyzed in the frequency domain [10].

To surface privacy concerns, Molina-Markham et al. recently highlighted that monitoring the home's aggregate power consumption can reveal private information about a homeowner's activity [21]. This work used methods from prior work in the energy monitoring and ubiquitous computing communities. Specifically, these authors use power usage patterns of the on/off states of electrical device to show that

privacy inferences are possible using seemingly innocuous data (e.g., power activity every few hours during the night may indicate regular feedings of a newborn).

Unlike prior work, which analyzes the state and identity of appliances, we use a different method to track a continuously varying signal. If we compare the EMI acquisition sensor to a microphone, an appropriate analogy would be that previous work has been able to detect the presence of speech and potentially who is speaking while our contribution enables us to understand *what* is being said. Concurrent to our own work, Clark et al. study powerline information leakage when a PC visits a Web page [5]. They evaluated their approach with eight websites and a single PC; their measurements instrument the connected outlet, whereas our measurements can take place anywhere on the powerline.

Security & Information Leakage. In the security literature, evidence suggests that the government has long known that ancillary EM emissions from CRT monitors leak private information about what those devices might be displaying [8, 12, 16]. Early work on studying electromagnetic leakage from CRTs has since been extended to flat-panel displays [17] and wired and wireless keyboards [3, 28].

There are important commonalities between all these works, e.g., they all seek to extract as much information as possible from the observed EMI noise. There are, however, also important differences. We focus on modern LCD and Plasma systems manufactured between 2007 and 2010. More fundamentally, we find that there are inherent differences in the types of signals we analyze. For example, both [3] and [28] leverage EMI resulting from a keyboard's digital transmission of bits corresponding to key presses. Past works on CRTs/LCDs similarly looked at analog/digital transmissions within the TVs [8, 12, 16, 17], however there is an *additional* level of indirection in the EMI information leakage that we exploit in this paper. Specifically, although the TV's video content is transmitted digitally within the TV, we do not knowingly extract any information about that digital content. Rather, we extract information from EMI emitted from the TV's power supply as it drives the video display, which consumes varying power.

The security community has also explored numerous other classes of information leakage, including: the time to perform various tasks (e.g., [13]), optical emanations (e.g., [19] for network appliances and [15] for CRTs), acoustic emanations (e.g., for printers [4], CPUs [26], and keyboards [1]), and reflections (e.g., [2]). There is also security research focused on the modern television showing that it is possible to infer what someone is watching over an encrypted wireless video stream from the size of the transmitted packets [25]; that approach exploits information leakage through variable bitrate encoding schemes, which was pioneered in [29]. Related to power consumption, but far afield from our own work, is the broad area of power analysis and differential power analysis for cryptographic processors [14].

4. CONTEXT AND THREAT MODEL

This work focuses on understanding and assessing how information might leak from a highly popular consumer electronic device: the TV. In this section we characterize the threat models under which our results are the most applicable. We stress, however, that (in our opinion) the results in this paper are of scientific interest independent of any particular threat model or policy question.

We consider three key axes for our threat model: the type of attacker, the adversarial goals, and the adversarial resources. We consider each in turn.

Attacker Type. We consider two key types of attackers:

- **Intentional Attacker.** An entity seeking to violate the privacy of an individual; may attach an adversarial monitoring device to the victim's powerline.
- **Unintentional "Attacker."** A party that attaches a monitoring device to a powerline for *legitimate* purposes (e.g., a smart meter or a ubicomp device). The word "attacker" is in quotes because the party in question is not malicious. Rather, the attached device will record measurements from the powerline for *non-malicious* purposes. The device may store those measurements locally or transmit (a function of) those measurements to an external server for processing and/or storage.

An intentional attacker could, for example, plug a sensor into an exterior outlet on a neighbor's home and thereby monitor EMI on the neighbor's powerline. We observe, however, that if the attacker could enter the neighbor's home, then the attacker might also explore even more devious attacks, such as hiding a webcam in a sensitive location. Thus, the most likely intentional attacker is one that does not have direct physical access to the monitored space (the interior of a home, apartment, or dorm room) but that does have direct physical access to the powerline (e.g., via an external outlet). The device could also be disguised as a different device, such as a harmless power adaptor or DVD player.

A more pressing concern is likely the unintentional "attacker." Consider, for example, this scenario recounted in a recent joint filing before the State of California by the Center for Democracy & Technology and the Electronic Frontier Foundation [20]:

> Government agents issued a subpoena to the suspect's utility to obtain energy usage records and then used a utility-prepared "guide for estimating appropriate power usage relative to square footage, type of heating and accessories, and the number of people who occupy the residence" to show that the suspect's power usage was "excessive" and thus "consistent with" a marijuana-growing operation.

As smart meters and ubicomp devices begin to collect increasingly fine-grained measurements from the home powerline for legitimate purposes (e.g., demand prediction, activity recognition, green energy recommendations), it is important to understand how those measurements could be misappropriated in the future.

Adversarial Goals. We consider two adversarial goals:

- **Target Video Identification.** An adversary might seek to determine whether the victim is watching a *target video.*
- **Target Channel Identification.** An adversary might seek to infer what TV channel the victim is watching.

In both cases the adversary might either be an intentional attacker or someone analyzing the data collected by an unintentional "attacker." For target video identification, the adversary may seek to determine whether the victim is viewing a video from a "watch list" of sensitive titles, e.g., politically

sensitive videos, religious videos, terrorist training videos. For target channel identification, we observe that the channel a person watches may reveal private information, e.g., political views, religion, known languages. The severity of these adversarial goals may be a function of the environment in which the victim lives. Within the U.S., we recall that an individual's video viewing records is considered highly sensitive pursuant to the Video Protection Privacy Act.

Our choice of these adversarial goals is driven by scale. For target channel identification, the attacker may know that a victim is watching one of a handful (a dozen or possibly a hundred) channels. Determining which channel a person is watching out of a small finite set of possible channels is far more feasible than determining which video a person is watching from the set of all possible videos. For target video identification, the goal is still significantly more feasible than determining which video a person is watching from the set of all possible videos — rather, the adversary must only be able to detect the viewing of specific videos.

Although we did not explore the following, we conjecture that our approach could also be used to estimate whether an individual is watching a pirated movie, *if* that pirated movie has some (possibly slight) deviation from the authorized copy of that movie. For example, if a pirated movie begins with a characteristic opening sequence — e.g., the opening animation inserted by the pirating organization, or an opening sequence that announces that the video is for the exclusive use by members of the Academy of Motion Picture Arts and Sciences — then that opening sequence might reveal itself in the leaked EMI. We explicitly did not investigate the identification of pirated content. But, as evidence of plausibility, over the course of our experiments we did develop the capability to visually distinguish between the splash screens for different studios (e.g., MGM, Twentieth Century Fox, etc.). The lack of these splash screens or the FBI warning screen on a DVD could also be an indicator.

Adversarial Resources. Finally, in addition to assuming that the attacker has the ability to obtain sufficient quality measurements from the victim's powerline, we assume that the attacker knows the type of TV that the victim is using. Although not the focus of this paper, we consider extensions in Section 10 where this is not necessary. Although outside the scope of this paper, we also conjecture that a second level classifier could be designed to infer the model of TV based on the operational characteristics of the SMPS, which is also discussed in prior work [10]. Finally, we assume that the attacker has access to the videos that he or she seeks to identify (e.g., DVDs for the target videos or feeds from all the channels that the victim might be watching).

5. SYSTEM DESCRIPTION

We now turn our attention to describing the data acquisition and post-processing components we used for our experiments. See Figure 2.

EMI Capture Hardware and Signal Processing. Our system consists of three main components. First, we connect a custom power line interface (PLI) module to any available electrical outlet in the recording environment to gather the conducted EMI signal. Second, we use a high-speed data acquisition module to digitize the incoming analog signals from the PLI. Lastly, we use a data collection and analysis PC running our software to condition and process the incoming signals from the digitizer.

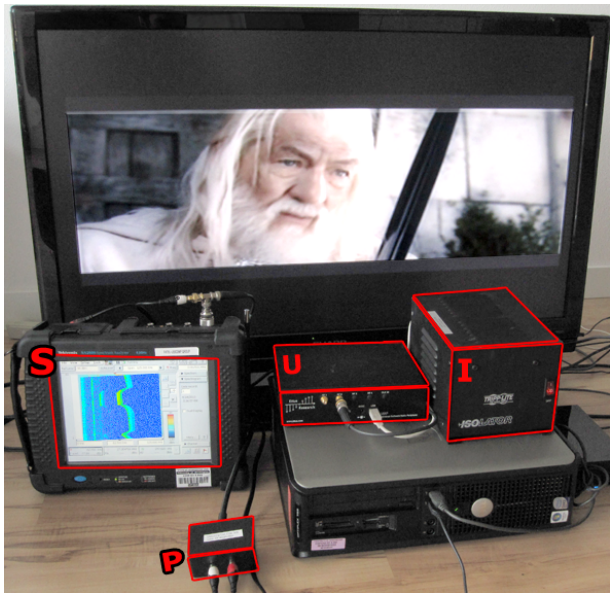


Figure 2: Components of our EMI capture hardware: Power Line Interface (P) for filtering 60 Hz signal voltage sensing, high speed digitizer (U) for analog to digital conversion of signals and an isolation transformer (I) for filtering EMI from other devices in our laboratory environment. The Sharp 42" LCD TV and the data logging PC are also visible. We also connect a spectrum analyzer (S) for debugging purposes and for visualizing the real-time EMI signal as a waterfall plot.

Harvesting EMI. The power line interface (PLI) that we designed to capture EMI was modeled on the circuit described in [10] with slight modifications to increase sensitivity at a broader frequency range. Our PLI module is essentially a high pass filter, composed of 3 resistors and 2 capacitors arranged as depicted in the schematic in Figure 3. This circuit design yields a RC high-pass filter which has two important properties. First, as in [10] it removes the 60 Hz frequency which is the rate at which alternating current (AC) is supplied in the U.S.; if this filter is not included, the 60 Hz signal would overwhelm our capture system yielding unusable data (and also potentially damaging any connected low voltage hardware). Second, the filter has a flat frequency response between 6 and 30,000 kHz which allows us to capture any conducted EMI in this expansive frequency band.

Post Processing and Automation. To transform the EMI data into the frequency domain, we fed the analog signal from the PLI into a USRP (Universal Software Radio Peripheral), which acts as a high-speed digitizer. We set the sampling rate of the USRP to 500 kHz, which allows us to effectively analyze the spectrum from 0 to 250 kHz (under the Nyquist Theorem). We found no interesting television signals beyond 250 kHz.

The digitized data from the USRP is then streamed in real time over a USB 2.0 connection to a PC running customized GNU Radio software. The PC processes the incoming data and computes a Blackman-Harris windowed (size half of FFT

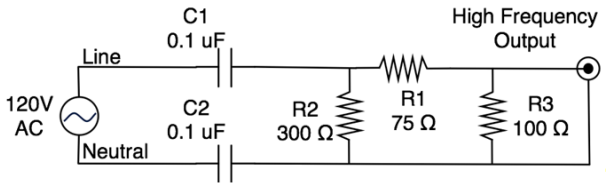


Figure 3: Schematic of the Power Line Interface, modified from [10] for broader sensitivity.

length) Fast Fourier Transform (FFT) 122 times a second to produce a vector of 2048 points which are spread uniformly over the spectral range (from 0 to 250 kHz). Next we compute the squared magnitude of each FFT vector and store the result on disk for post-processing. Our experimental infrastructure is highly automated; it can automatically stream video content to TVs while simultaneously logging EMI signals.

Lab vs Home Setup. Over the course of our experimental analysis we collected EMI data from lab and home settings. During lab recordings, we connected the TV being tested and our PLI sensor to the output of an isolation transformer (model Tripp Lite IS500HG). The isolation transformer reduces electrical noise and EMI present on the power line and presents a cleaner AC power at it's output [2]. Generally, isolation transformers are used to protect against electric shock, or to suppress electrical noise in sensitive devices such as high-end audio systems.

Conversely, in home recordings, we did not use the isolation transformer but rather connected the PLI directly to any available electrical outlet. This meant that we were no longer guaranteed a clean electrical background, and our target EMI could now be affected by the noise produced by any of the devices on the powerline.

6. SELECTION OF TVS AND MOVIES

To evaluate the level of EMI based information leakage in modern TVs, we procured a set of eight TVs that differ along a number of axes including: manufacturer (Panasonic, Samsung, Sharp), display technology (LCD or Plasma), size ($32''$, $42''$, $58''$), and date of manufacturing (2007–2010).

Table 1 summarizes our hardware selection. We opted for an equal distribution of LCDs and Plasmas — the two dominant technologies currently on the market. Furthermore we deliberately selected three pairs of duplicate TV models to enable analysis of EMI signature similarity across identical hardware. This table also introduces a naming scheme which we use in the remainder of the paper. Under this convention, television names include manufacturer, size, and model-pair information. For example the name Samsung-32 indicates that the Samsung TV is $32''$ and that we have only one model instance; similarly, Panasonic-42-A is a $42''$ Panasonic TV which is the first of a pair (as indicated by the "A").

Next we needed to choose what video signals to send to the TVs while recording EMI. For experimental feasibility, we opted to create a 20 movie database. To make our movie selection systematic, we chose four top-rated films within each of five distinct genres (ratings were extracted from the Internet movie database, `imdb.com`). Since these movies had variable running times, we limited our analysis to the first 60 minutes of each DVD video stream. Table 2 lists the movies that constitute our library. Looking ahead, our results suggest that our methodology is applicable far beyond a collection of 20 target movies. Nevertheless, we note that past work on video information leakage also studied a similarly-sized collection of movies (26 movies), though in a very different context (information leakage from streaming encrypted video data) [25]. Returning to the threat model in Section 4, we also argue that distinguishing between 20 movies approximates the adversarial goal of target channel identification. In our experiments we also embed random portions of these 20 movies into streaming TV signals and assess our ability to detect the embedded videos, thereby approximating the adversarial goal of target video identification.

[2]We recognize that a 'Line Impedance Stabilization Network' LISN device would have provided a more optimally stable and repeatable signal; however, we expect a LISN to only improve the already high matching accuracies achieved with our acquisition hardware.

Genre	Movie
Action	(1) Lord of the Rings: Return of the King, (2) Star Wars V: Empire Strikes Back, (3) The Bourne Ultimatum, (4) The Matrix
Animation	(5) Wall-E, (6) Shrek 2, (7) The Lion King, (8) Aladdin
Comedy	(9) Office Space, (10) Meet the Parents, (11) The Hangover, (12) Wedding Crashers
Documentary	(13) Planet Earth: Fresh Waters, (14) Food Inc., (15) An Inconvenient Truth, (16) Top Gear (Season 4;Episode 7)
Drama	(17) The Shawshank Redemption, (18) American Beauty, (19) Titanic, (20) Requiem for a Dream

Table 2: List of twenty movies that we included in our dataset selected such that they span various genres.

7. DATA COLLECTION PROCEDURE

We performed recordings on each of our 8 TVs in a lab environment where we could carefully regulate the electrical conditions. This data accounted for 13 days of EMI traces, and provided the foundation of most of our analysis. Next, we selected a single TV (Panasonic-42-B) and deployed our system to three home locations to gather data in electrically uncontrolled settings. Lastly, we recorded EMI from 20 hours of broadcast programming, which was the basis of the experiment described in Section 9.5.

Lab Recording. During lab recordings, we plugged a given TV and our voltage sensor into an isolation transformer to remove interference from other electrical devices in the building. We then streamed our database of 20 movies twice[3] while logging EMI in the 6 to 250 kHz range.

Home Recording. To evaluate EMI based information leakage under a more naturalistic setting, we also performed data collection in three homes with distinct sizes, neighborhoods, ages (1906, 2003, 2009), and styles (apartment, suburban house, multi-family home) on a subset of our movie database (12 movies, 15 minutes each). Table 3 shows the

[3]Two runs of our movie library were recorded on each TV to investigate the repeatability of EMI signatures.

ID	Style/ Year Built	Size/Floors
H1	Single family Home/2003	3000 sq. ft/ 2 flrs
H2	Apartment/2009	657 sq. ft/ 1 flr
H3	Multi-family Home/1906	800 sq. ft/ 3 flrs(*)

Table 3: A summary of the homes showing the style, size and number of floors. (*) 800 sq. ft refers to only one apartment in a converted multi-family home.

specifications for these residences. During the home data collection, we did not prevent occupants from actuating electrical devices and we also did not use the isolation transformer. Also, we plugged in our sensor in a arbitrarily chosen electrical outlet that was available. Thus, we captured not only the EMI from the TV but also any number of active electrical devices operating in the home under natural conditions (e.g., lamps, kitchen appliances, computers, other televisions). Many times during our home data collection EMI from other devices overlapped with the EMI from the TV under investigation. In addition, Home#2 used power line communication (PLC) during recordings, however we did not observe significant signal interference since home PLC signals are typically narrow-band in our recording range.

8. SIGNAL CONDITIONING

Automated Signal Identification. To facilitate analysis we needed to devise a method for extracting meaningful signals from the FFT data captured using our hardware. Recalling the properties of EMI we described in Section 2, we see that there are two critical questions which must be answered in order to track the signature of a particular device: (1) in what range of the frequency spectrum does the EMI signal reside, and (2) in what way does the EMI signal fluctuate with changes in device state.

Although it is often possible to answer both of these questions by visual inspection of a spectrum analyzer during device operation, we sought to develop a systematic search procedure which played a 5 minute custom software-generated video while performing a range sweep to find EMI changes that are strongly correlated in time with the controlled signal. The search method was designed to determine the EMI frequency range, fluctuation type (frequency or energy density), and to ultimately produce a 1-dimensional time series representation (EMI trace) normalized to the 0:1 domain.

The search algorithm requires FFT samples extracted from EMI over the 0 to 250 kHz frequency range collected during controlled video signal playback (36600 total FFT samples). Given this input our procedure analyzes subsets (windows) of the frequency spectrum and correlates their EMI trace (computed as both energy density and frequency shift) to the trace of brightness changes in the software-generated video [4]. At the end of the sweep, the optimal 1-dimensional EMI trace is returned which produced maximal correlations to the control signal (the search is performed across all frequency windows and signal types). The functions we applied are shown below in pseudo-code using the array syntax *initial-value:final-value* common to the MATLAB programming language. Smoothing is done using a 4th order Butterworth filter with a cutoff frequency of 10 kHz; decimation is done using eighth-order lowpass Chebyshev Type I filter with a cutoff frequency of $0.8 * (Fs/2)/r$ where $r = 20$ is the decimation factor; standard normalization is performed to shift the domain to 0:1.

[4]Our experiments led us to conclude that artificially generated color gradients work well as control signals.

Television Name	Freq. Range	Signal Type
Panasonic-42-A/B	1 to 60 kHz	Energy Density
Samsung-58-A/B	45 to 55 kHz	Frequency Shift
Samsung-32	10 to 50 kHz	Energy Density
Sharp-42	60 to 90 kHz	Frequency Shift
Sharp-32-A/B	35 to 40 kHz	Energy Density

Table 4: Frequency regions and EMI signal types for our 8 TVs.

Using the EMI search procedure, we were able to characterize the EMI signatures for all of the TVs in the lab setting (Table 4). In addition, the search method proved invaluable when we analyzed EMI from home recordings as different residences had unique levels of dynamic background noise which overlapped with tracking ranges we found in lab; our methods were always able to find frequency regions minimally influenced by the other powerline noise and thus enabled us to recover the TV EMI signal (due to harmonics).

EMI search procedure:

```
var EMI := fft(range := 0 : 250kHz, time := 1 : 36600);
  var target := brightnessGradient(controlVideo);
    begin
      for r := 1 to 250 step 2 do
          for len := 10 to 60 step 10 do
              if r + len > 250 then continue fi;
              rEnd := r + len;
              [freqEMI, densEMI] := extractEMI(r, rEnd);
              fCorr(r, len) := xcorr(freqEMI, target);
              dCorr(r, len) := xcorr(densEMI, target);
          od
      od
      [optRange, sigType] := max_{r,len}(fCorr, dCorr)
      proc extractEMI(start, end) ≡
        type1 := sum(EMI(start : end, 1 : 36600);
        type2 := maxIndex(EMI(start : end, 1 : 36600);
        densEMI := normalize(decimate(smooth(type1)))
        freqEMI := normalize(decimate(smooth(type2)))
      .
    end
```

9. ANALYSIS AND RESULTS

We began by experimentally establishing that individual televisions produce repeatable EMI traces when the same video content is shown multiple times. Building on this result, we show that the EMI signals produced by multiple TVs of the same model are highly correlated given identical video inputs. We then use our collection of 20 movies to study the ability of an attacker to match the EMI collected from one movie with a database of previously collected EMI measurements from all 20 movies. Together, these results suggest that an attacker can use TV EMI to accomplish the target channel identification goal from Section 4 with high confidence. Next, we study the feasibility of our approach in three home environments, where additional signals sources are active on the power line. The data from homes also allows us to reason about the possibility of matching EMI traces gathered in an electrically clean setting to those collected

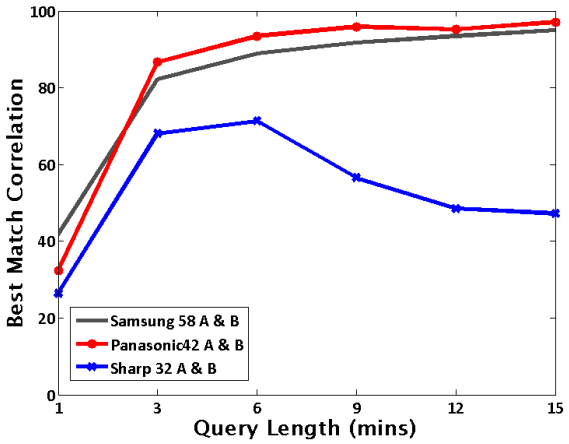


Figure 4: Cross Correlation results at various query lengths.

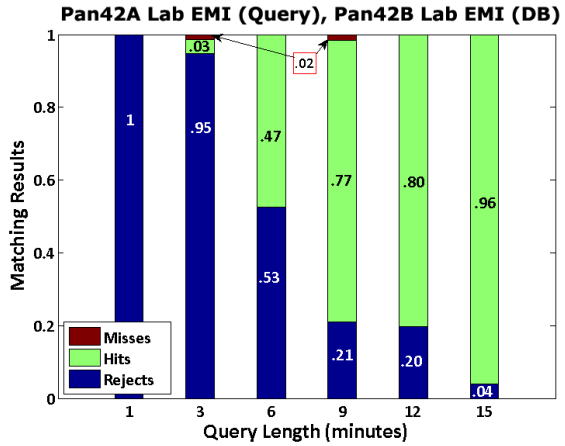


Figure 5: Matching results for Pan-42-A vs Pan-42-B for varying query lengths.

from noisy environments (i.e., matches are possible between lab and home EMI). Finally, we showcase the robustness of our methods by accurately identifying segments of EMI from our movie database that have been embedded within a stream of foreign EMI (20 hours of local television broadcasting), thus providing strong evidence that an attacker can also use TV EMI to accomplish the target video identification goal from Section 4 with high confidence.

9.1 Signal Repeatability in Lab

Individual TVs. The first question we sought to answer was whether repeated video content played on a target TV would produce consistent EMI. To test this, we computed the cross-correlation of a movie's EMI trace between multiple recording sessions. The left column of Table 5 contains the cross-correlation statistics when this computation was applied to every movie in our database and repeated for each TV, using only the first 15 minutes of each trace. These levels of high cross-correlation suggest that for most televisions multiple runs of identical video content produced consistently similar EMI. The exceptions, which we consider in more

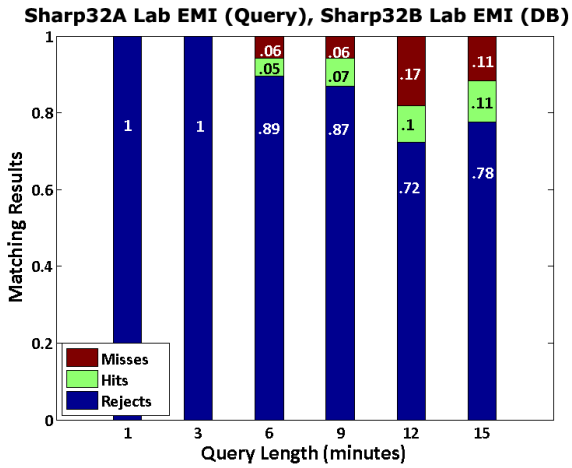


Figure 6: Matching results for Sharp-32-A vs Sharp-32-B for varying query lengths.

depth later, are the Sharp 32″ TVs. Conversely, dissimilar video content tended to produce highly uncorrelated EMI traces. The latter conclusion was drawn from the results of our followup experiment which analyzed the potential for similarity between EMI traces from different video streams. More specifically, we recorded the highest levels of cross-correlation possible while matching a movie against all other films in the database (itself excluded); our measurements are reported in the right column of Table 5.

Model Families. While the above findings indicate significant information leakage within a TV, it's not clear whether this leakage poses a significant threat to privacy. The question arises: can an attacker infer information from the EMI of a TV that he or she is not able to physically access? We explore that question here by assessing how similar the EMI is between two TVs of the same model. We performed this test by replicating the methods above, but instead of comparing EMI signals from multiple recording sessions on a single TV we analyzed the correlations between EMI signals between single sessions of EMI recorded on TVs within the same model family (i.e., Panasonic-42-A vs Panasonic-42-B, Samsung-58-A vs Samsung-58-B, and Sharp-32-A vs Sharp-32-B). As before, we computed the EMI similarity of identical content (Table 6, left column) as well as distinct content (each movie against all others, Table 6, right column).

Discussion. The experiments suggested that, with few exceptions, EMI from identical video signals are highly repeatable while distinct video sources yield independent EMI signals. These results were shown to hold within individual TVs as well as across pairs of TVs within a model family. In both cases, the Sharp 32″ TVs produced outlier data points, reflecting a significant decrease in EMI consistency relative to the averages of other TVs. We attribute this to the fact that these devices had the least dynamic EMI (minimally modulated by screen content changes) within the televisions we tested. We tested the power consumption of the Sharp 32″ TVs and found that their power draw changed by only 1W (175W to 176W) when playing a software generated custom video, whereas the Sharp 42″ TV exhibited a 24W (190W to 214W) change using the same video signal. Thus we attribute the small dynamic shifts in EMI of the Sharp 32″ to

TV	Cross Correlation Same Content (%)		Cross Correlation Different Content (%)	
	Average	Standard Deviation	Average	Standard Deviation
Panasonic-42-A	98.18	±2.77	60.36	±12.83
Panasonic-42-B	98.99	±0.98	59.51	±12.97
Samsung-58-A	97.47	±1.97	53.29	±10.22
Samsung-58-B	96.71	±2.73	53.37	±9.70
Samsung-32	98.26	±5.21	65.56	±16
Sharp-42	97.30	±8.26	63.45	±12.84
Sharp-32-A	60.03	±8.12	56.41	±5.10
Sharp-32-B	60.91	±8.39	56.39	±5.79

Table 5: Analysis of EMI signal similarity given identical and different video content within a TV - averaged over the 15 minute intro segment of all movies.

TV	Best Cross Correlation Same Content (%)		Best Cross Correlation Different Content (%)	
	Average	Standard Deviation	Average	Standard Deviation
Panasonic-42-A and B	96.82	±4.67	59.93	±12.85
Samsung-58-A and B	94.96	±9.23	59.30	±10.09
Sharp-32-A and B	77.21	±18.84	53.31	±12.96

Table 6: Analysis of EMI signal similarity given identical and different video content between TV pairs (averaged over all movies). For this we used randomly selected 15 minute streams from each 60 minute movie

the power consumption characteristics of the circuitry. There is also a possibility that this particular model incorporates a fixed frequency power supply.

9.2 Signal Matching and Query Length

Our initial experiments suggested that extracted EMI signals are consistent within TVs as well as between identical models using cross correlations computed over 15 minute segments of content. We next evaluated how varying the length of EMI traces impacts signal repeatability and differentiability.

We designed a search procedure that would take as input a full EMI trace of a movie and extract 10 query segments (of variable lengths) from multiple (randomly chosen) starting indexes along the 60 minute stream. Each query segment would then be matched against the EMI signals of the entire movie database (using sliding-window cross correlation), and the highest correlated match would be returned. This technique is also referred to as matched filtering in signal processing communities. We applied this procedure to find matching signals between TVs of the same model; query data would be extracted from the EMI library of the first TV in the pair, and the database signals would come from the other (e.g., query = EMI from Panasonic-42-A, database = EMI from Panasonic-42-B).

We performed a series of runs using query lengths ranging between 1 and 15 minutes (1, 3, 6, 9, 12 and 15 mins) and in each instance computed the cross correlation between the query and its best match (averaged across 10 samples for each query length). The matching results for the three TV model pairs are show in Figure 4. Even short length queries tend to find highly correlated matches in the database. In particular, for TVs with repeatable EMI signals, once the query length reaches 6 minutes the correlation reaches 93.7%.

Interestingly, for the Sharp-32-A and B TVs, longer queries led to degraded performance. We attribute this to the weak EMI signatures of these TVs (as previously mentioned) which are susceptible to noise whenever the EMI signal is not being modulated along its entire dynamic range. However, the Sharp-42 performed very well with just short queries.

9.3 Reducing False Matches

Usually the best correlated match to a query (returned by our search procedure) was taken from the trace of the same movie in the database; however this was not always the case and we address this issue in the current section. To develop a method that would reduce false matches, we modified our matching algorithm to withhold query responses unless the top correlation match was a "clear winner" and was separated from the next best candidate by a margin of 5%. We experimented with multiple settings of this threshold parameter (including setting its value to be proportional to the number of samples in the query) but found that the 5% level provided a good trade-off between numbers of false positives and rejected samples given the query lengths and dataset size we tested (we provide a lengthier discussion of the threshold value choice in the Appendix B).

Given this modified search procedure, each query could be classified into one of three bins:

- **Miss:** the search engine is confident in the match (accept) but there is a mismatch between the search engine's best guess and the query origin. (Failure.)
- **Hit:** the search engine is confident in the match (accept) and the match was the movie from which the query itself was extracted. (Success.)
- **Reject**: the best match was a not a "clear winner," and the matching algorithm chooses not to respond. (Neither success nor failure.)

In Figure 5 and Figure 6 we respectively show the distribution of queries into these bins as a function of query length for the TVs which were the best (Panasonic 42″) and worst performers (Sharp 32″) on the test for signal repeatability within a model family.

We note that the Sharp 32″ A vs B data causes a very high rejection rate due to the consistent lack of a "clear winner,"

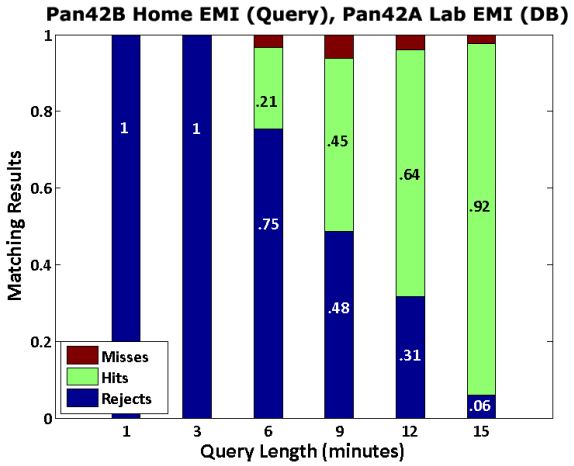


Figure 7: Distribution of queries into rejected, hit and miss for lab vs home data for Panasonic-42.

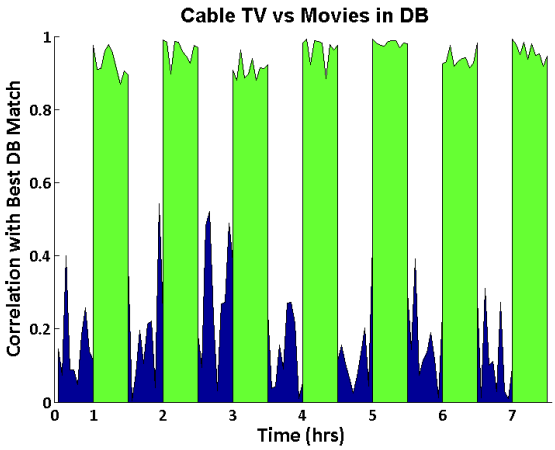


Figure 8: Cross correlation for target (green) movies embedded in broadcast content (blue).

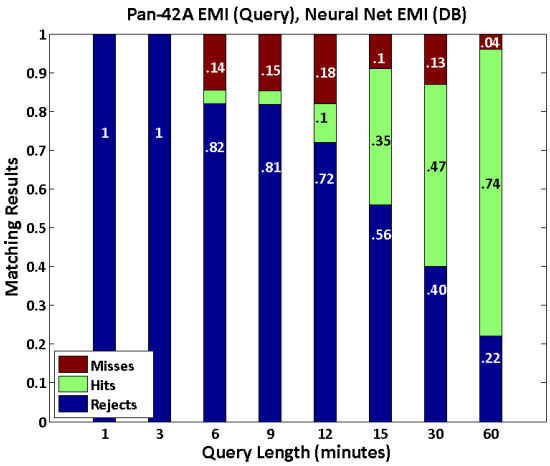


Figure 9: Matching results for Panasonic-42-A EMI queries against neural network database.

in addition further strengthening the hypothesis that the Sharp 32″ A & B are very sensitive to noise. Conversely the Panasonic 42″ A vs B data quickly reaches a high hit rate (and low rejection rate). In all cases the miss rate was low. This method is therefore less influenced by noise; queries that result in a match (non-reject) have a high probability of being correct.

9.4 Home Results

Having found evidence of significant information leakage within a lab setting, we next turned to evaluating our approach in a more natural environment. We wanted to deploy a TV in multiple homes and see if we could match queries from home EMI to lab EMI datasets. In choosing which television to utilize for our residential recordings, we initially restricted the candidate pool to TVs for which we had duplicate models (Panasonic-42-A&B, Samsung-58-A&B, and Sharp-32-A&B). We rejected the Sharp 32″ TVs based on their weak EMI signatures. Given the choice between the Panasonic 42″ TVs and Samsung 58″ TVs, we selected the smaller TVs because they were significantly easier to transport.

We set up our system (Panasonic-42-B, PLI, and logging equipment) in three different homes (see Table 3, along with Appendix C) and in each context recorded a smaller version of our database (3 hours total — first 15 minutes of the first 12 movies). Next we matched queries extracted from home EMI (Panasonic-42-B) against the EMI signature database collected in the lab (Panasonic-42-A) using the search system which had the capacity to reject searches that did not produce "clear winners" (as described previously). The results of this analysis are shown in Figure 7.The hit rate for 15 minute queries drops from 96% in the lab environment with no misses to 92% in the home environment with a 2% miss rate.

9.5 Searching for Target Content

The above results show that our small collection of 20 movies yield distinguishable EMI when displayed on modern TVs; this emulates the target channel identification adversarial goal in Section 4. We now turn to experimentally emulating a target video identification attack scenario (recall again Section 4). Under this scenario, a person might be watching arbitrary TV content, and the attacker wishes to determine whether that content corresponds to some target (e.g., censored, banned, sensitive) content. To explore this adversarial goal, we recorded 20 hours of EMI while the TV was tuned to over-the-air HD television broadcast.

Figure 8 shows the results of running our matching framework on EMI from one hour long chunks of over the air broadcast interleaved with one hour movies (the targets) from our database. Given a query length of 6 or more minutes the over the air cable data never returned a match because matches could not pass the confidence threshold. Conversely, the content embedded from our "sensitive" database had a rejection rate of 3.3% and a hit rate of 95.06%. This demonstrates that target videos were clearly detectable when intermingled with non-target videos.

10. EXTENSION: MODELS OF EMI

The results in the previous sections motivated us to ask the following question: can we reverse engineer the method by which a TV produces EMI as a function of its video input? Such a tool could be used to build a database of EMI models to predict noise signatures without needing physical access

to target devices (after the training phase). It would also be plausible to learn to recreate EMI from a home TV without ever seeing a TV of that type, provided that there are samples of data from periods during which the TV displays known content (e.g., if the user is watching one of five news channels in the morning, the system could build a model for the TV to try and detect whether the user is watching some specific show later).

We investigate the plausibility of learning a model to recreate the EMI of one TV (Panasonic-42-A) by framing the problem as an instance of supervised learning, where the goal is to approximate how the various hardware components of the device (taken together as a black box) modulate the SMPS and produce EMI from video input. To tackle this problem we chose to use recurrent neural networks (RNN) because the EMI we are seeking to replicate is a continuous time signal which exhibits non-stationarity (i.e., the current state of the EMI depends on the past several samples) hence ruling out the possibility of using popular stationary methods for supervised learning (i.e., Support Vector Machines). A further motivation to use RNNs comes from their ability to accommodate for non-linear interactions between the input and output which we could not rule out from the processes that shape electromagnetic interference.

Features from Video Frames. The input to our model, is a sequence of video frames arriving at a rate of 30FPS (NTSC) with typical DVD resolution of 720 x 480 pixels. In its raw form the input dimensionality at each sample is extremely high (3 x 720 x 480 per frame) and prohibitively large for model training. Thus, we opted to compress each video frame into a 11 element vector which extracts selected features (meant to capture key statistics for each frame) from the visual content and greatly reduce the complexity of the learning problem. The features we compute from each frame are listed below:

- **Brightness:** cumulative sum of averaged RGB intensities (based on pixel values).
- **Flux:** change in brightness between consecutive frames.
- **Edge Intensity:** pixelsum of a Canny Edge filter.
- **FFT:** slope of the best fit line to an FFT.
- **Color:** mean and standard deviations for gaussians fitted to R, G, and B color histograms (6 params).
- **Bitrate:** kbits/second computed using FFMPEG.

These features were post processed in a fashion very similar to the signal conditioning of the EMI. In particular, we performed normalization (0 to 1 scale), smoothing (Butterworth filter with 10kHz cutoff frequency) and decimation (10 samples per second).

Network Structure. RNNs are a class of neural networks in which intermediate layers (i.e., those separating input and output) have connections to neighboring layers as well as (re)connections to themselves; these properties lead to self feedback (memory) which enable dynamic temporal behavior [20]. At time t the network input layer consisted of a video frame represented as a 11 element feature vector. The input layer was connected to the first of 3 hidden layers (connected in succession, each composed of 11 neurons to match the dimensionality of the input) and the final hidden layer was connected to a scalar output layer representing the normalized time series EMI at time t.[5]

Training. The training phase began with randomly initialized network parameters which were tuned using backpropagation through time (BPTT) via the Levenberg-Marquardt gradient method. The criterion for performance was how well the network output matched test EMI (measured as mean squared normalized error). Each training session concluded when the optimization converged or after 50 epochs (whichever came first).

Results. We used our neural network to generate a database of *synthetic EMI* (i.e., given *only* video data) and performed matching searches using *real* TV EMI traces as queries. Prior to defining our training and target sets, we systematically removed the five movies with the globally lowest self similarity measures across all TVs (Wedding Crashers, The Hangover, Meet the Parents, Office Space, Top Gear). This eliminates noise artifacts during training *and* approximates the goal of only plausibly detecting target videos with high self similarity.

The remaining 15 movies were divided into random sets of 10 training movies and 5 target movies, and for each training/target combination we evaluated matching performance using our thresholding method (Section 9.3) which rejected queries unless there was a .05 gap in the cross-correlation of the best candidate match and the runner up. For each combination of training and target movie sets we computed matching using a database of neural network EMI (generated for the 5 target movies) and queried using real TV EMI (from 10 test movies — the 5 target movies with generated EMI via the neural network and the 5 movies originally excluded due to low self similarity). The average cross correlation of the neural network output and the ground truth EMI of the Panasonic-42-A for the 5 target movies was .88 (for a query length of 60 minutes). Figure 9 shows the results of submitting real EMI for the 5 target movies into our query matching system (the hit rate at 60 minutes is .74, lower than .88, because of the .05 threshold gap). Figure 10 visually shows the alignment between the generated EMI for a target movie and real EMI for that movie. The matching database correctly rejected EMI from the five movies originally excluded because of low self similarity with average probability 96.8 (averaged across query lengths longer than 15 minutes).

Using brightness as an isolated input to train a network generates signals that are .73 correlated with real EMI for the target movies (other features produce an average of .17 correlation). In future work it would be worthwhile to use hypothesis driven regression models to strengthen our understanding of generative EMI sources, however such analysis will require careful experimentation with secondary sources of EMI signal modulation (e.g., display processing circuitry) â as suggested by our results and confirmed by [18].

It is important to stress that the above analysis should only be viewed as evidence of plausibility — the primary focus of our work is on the experimental analysis of real EMI as discussed in the previous sections. Our results here do, however, suggest that it is feasible to use supervised methods to train generative models of EMI.

[5]Though we cannot claim that this network structure is globally optimal for our problem, it offered high performance and low overfitting after experimenting with a large set of multi-layered neuron arrangements.

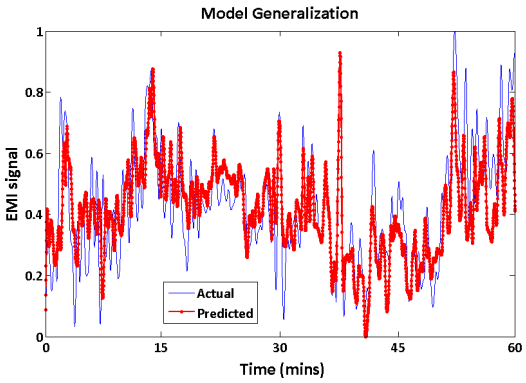


Figure 10: Neural network predictions (driven via visual features) vs. actual EMI of a 60 minute movie not included in the training set (Bourne Ultimatum).

11. DISCUSSION

Applicability to Other Devices. To better understand the scope of information leakage exposed via conducted EMI from SMPS, we executed a survey of other common home electronics: a personal computer, a laptop, a laser printer, a DVD player, a game console, and a washing machine. Our aim was to assess whether these SMPS-based appliances similarly produce noise signatures as a function of their device operation. Recall that past work on conducted EMI information leakage, such as works focused on keyboards [3, 28], observed EMI that was a direct function of electrical signals — bits — already communicated within the device in question; SMPS were not designed to carry digital signals, yet we observed that many appliances do yield EMI signal fluctuations that are correlated with the various functions supported by the device in question.

We found that the laser printer generated increased electromagnetic interference while warming up and during the printing process that engages the motors. In the case of the personal computer and laptop, we observed increases in EMI amplitude while running a collection of intensive CPU/memory/disk benchmarks. The DVD player's (with internal FM Radio) EMI was affected by volume changes, and the game console produced different EMI signals depending on whether it was in the main menu or in the process of rendering a game. The washer produced EMI bursts whenever the barrel rotated during the wash cycle.

Defenses and Challenges. There are several potential countermeasures for minimizing information leakage through EMI. One simple defense mechanism is to connect SMPS based devices to a powerline isolator similar to the one used in our laboratory experiments. The internal transformer provides enough isolation such that the high frequency noise does not pass back over the powerline (assuming the isolator itself has not been compromised). This capability is found in some newer line conditioners and in some instances of uninterruptible power supplies (UPSs).

A potential whole home solution, which does not require installing a device behind every electronic appliance, would be to inject random high-energy broadband noise over the powerline. In addition to traditional challenges with randomness, a key pragmatic challenge with this approach is that it must conform to FCC regulations. In addition, this would cause problems with legitimate powerline-based communication systems like broadband over powerline and X10(r) home automated systems. A more practical approach could be to identify potential devices that may be leaking information by observing the power line and only blocking certain frequency bands using an active noise rejection system or informing the home owner where to install line isolators.

It might be tempting to fall back on an out-of-band solution and ask for new regulation on how SMPS power supplies are built. In addition the political challenges with such a solution, this approach is also faced with the following tension: it may be impossible to fully defend against such information disclosure while still being in compliance with Energy STAR. Said another way, existing regulation within the U.S. may make it difficult or infeasible to implement defenses since the costs of privacy (increased consumption and decreased efficiency) are in direct conflict with recent legislation. This creates a fundamental challenge: to develop systems that are both highly energy efficient and private.

Through this contribution we assess the key features of SMPS information leakage over the powerline, and provide a knowledge base to aid in future efforts for mitigating the tension between device functionality and privacy.

12. CONCLUSIONS

We conducted an extensive study of powerline information leakage from eight modern TVs spanning two technologies (LCD, Plasma), three sizes (32″, 42″, 58″), and three manufacturers (Panasonic, Samsung, Sharp). We find that these TVs generally produce stable, robust EMI signals on the powerline that correlate with the video being displayed. We discuss two adversarial goals that do not require an adversary to pre-train on the set of *all possible* movies, but rather train on a *small set* of video content/movies (e.g., to infer whether a person is watching a target, sensitive video or whether a person is watching a target, sensitive channel), and we describe two different classes of adversaries (intentional adversaries that might attach malicious hardware to a home's powerline and unintentional adversaries — e.g., ubicomp devices — that might unwittingly collect and storing privacy-compromising measurements from the home's powerline). Next, we experimentally evaluate our approach with 20 movies in both a clean laboratory setting and electrically noisy home environments with other active devices connected to the powerline. Subsequently, we describe extensions to our core work that allow an attacker to infer information from a home's TV's EMI without ever requiring access to that TV or another TV of the same model (by dynamically learning a model of the TV's EMI signatures as a function of known video content). Lastly, we close with a discussion of some defenses, as well as fundamental challenges to fully defending against this threat (e.g., complete defenses may be in violation of Energy STAR).

Acknowledgements

We thank Lee Tien for his insights, feedback, and encouragement, and we thank a Shane Clark, Alexei Czeskis, Tamara Denning, Daniel Halperin, Eric Larson, and Franziska Roesner for valuable comments on earlier drafts of this paper. This work was supported in part by NSF Award CNS-0846065 and an Alfred P. Sloan Research Fellowship.

References

[1] Dmitri Asonov and Rakesh Agrawal. Keyboard acoustic emanations. In *IEEE Symposium on Security and Privacy*, 2004.

[2] M. Backes, M. Dürmuth, , and D. Unruh. Compromising reflections – or – how to read lcd monitors around the corner. In *IEEE Symposium on Security and Privacy*, 2008.

[3] Andrea Barisani and Daniele Bianco. Sniffing keystrokes with lasers/voltmeters: Side channel attacks using optical sampling of mechanical energy and power line leakage. In *Black Hat Technical Security Conference: USA*, 2009.

[4] R. Briol. Emanations: How to keep your data confidential. In *Symposium on Electromagnetic Security for Information Protection*, 1991.

[5] Shane S. Clark, Jacob Sorber, Kevin Fu, and Erik Learned-Miller. Current Events: Compromising Web Privacy by Tapping the Electrical Outlet. Technical Report UM-CS-2011-030, Department of Computer Science, University of Massachusetts Amherst, Amherst, MA, July 2011.

[6] G. Cohn, S. Gupta, J. Froehlich, E. Larson, and S. Patel. GasSense: Appliance-level, single-point sensing of gas activity in the home. In *International Conference on Pervasive Computing*, 2010.

[7] Analog Devices. Analog devices power measure ic for smart meters. Analog.com, `http://www.analog.com/en/analog-to-digital-converters/energy-measurement/products/index.html`, 2010.

[8] V.W. Eck. Electromagnetic radiation from video display units: An eavesdropping risk? *Computers and Security*, 4:269–286, 1985.

[9] J. Froehlich, E. Larson, T. Campbell, C. Haggerty, J. Fogarty, and S.N. Patel. HydroSense: Infrastructure-mediated single-point sensing of whole-home water activity. In *International Conference on Ubiquitous Computing*, 2009.

[10] S. Gupta, S.N. Patel, and M.S. Reynolds. ElectriSense: Single-point sensing using EMI for electrical event detection and classification in the home. In *International Conference on Ubiquitous Computing*, 2010.

[11] G. Hart. Advances in nonintrusive appliance load monitoring. In *EPRI Information and Automation Conference*, 1991.

[12] History of Tempest. Cryptome, `http://cryptome.org/tempest-old.htm`, 2000.

[13] P.C. Kocher. Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems. In *Advances in Cryptology*, 1996.

[14] P.C. Kocher, J. Jaffe, and B. Jun. Differential power analysis. In *Advances in Cryptology*, 1999.

[15] M.G. Kuhn. Optical time-domain eavesdropping risks of CRT displays. In *IEEE Symposium on Security and Privacy*, 2002.

[16] M.G. Kuhn. Compromising emanations: Eavesdropping risks of computer displays. University of Cambridge Technical Report UCAM-CL-TR-577, 2003.

[17] M.G. Kuhn. Electromagnetic eavesdropping risks of flat-panel displays. In *Workshop on Privacy Enhancing Technologies*, 2004.

[18] M.G. Kuhn. Compromising emanations of lcd tv sets. In *IEEE International Symposium on Electromagnetic Compatibility - Aug 11-19 (to appear)*, 2011.

[19] J. Loughry and D. A. Umphress. Information leakage from optical emanations. *ACM Transactions on Information and System Security*, 5:262–289, August 2002.

[20] Jennifer Lynch and Lee Tien. Joint comments of the Center for Democracy & Technology and the Electronic Frontier Foundation on proposed policies and findings pertaining to the smart grid. Comments to the the Public Utilities Commission of the State of California, `http://www.eff.org/files/CDTEFFJointComment030910.pdf`, 2010.

[21] Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin. Private memoirs of a smart meter. In *ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, 2010.

[22] S.N. Patel, M.S. Reynolds, and G.D. Abowd. Detecting human movement by differential air pressure sensing in HVAC system ductwork: An exploration in infrastructure mediated sensing. In *International Conference on Pervasive Computing*, 2008.

[23] S.N. Patel, T. Robertson, J.A. Kientz, M.S. Reynolds, and G.D. Abowd. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line. In *International Conference on Ubiquitous Computing*, 2007.

[24] Alfredo Rial and George Danezis. Privacy-preserving smart metering. Microsoft Research Technical Report MSR-TR-2010-150, 2010.

[25] T.S. Saponas, J. Lester, C. Hartung, S. Agarwal, and T. Kohno. Devices that tell on you: Privacy trends in consumer ubiquitous computing. In *USENIX Security Symposium*, 2007.

[26] M. R. Spiegel. *Theory and Problems of Probability and Statistics*, chapter 14 (Correlation Theory). McGraw-Hill, 1992.

[27] The benefits of smart meters. California Public Utilities Commision, `http://www.cpuc.ca.gov/PUC/energy/Demand+Response/benefits.htm`, 2010.

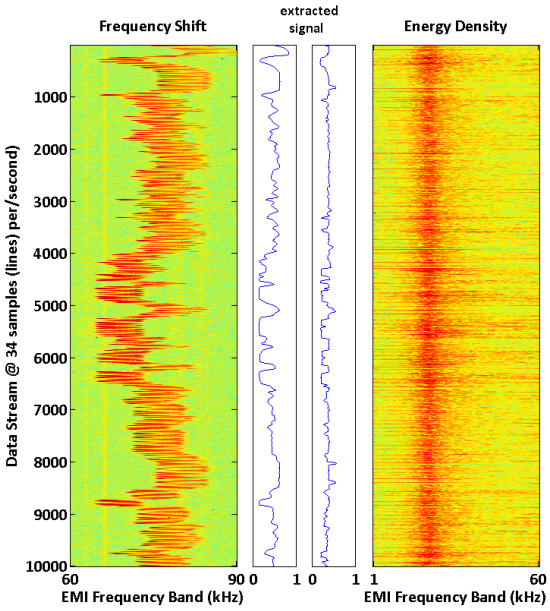
[28] M. Vuoagnoux and S. Pasini. Compromising electromagnetic emanations of wired and wireless keyboards. In *USENIX Security Symposium*, 2009.

[29] C. V. Wright, L. Ballard, F. Monrose, and G. M. Masson. Language identification of encrypted VoIP traffic: Alejandra y Roberto or Alice and Bob? In *USENIX Security Symposium*, 2007.

APPENDIX

A. TWO EMI TYPES

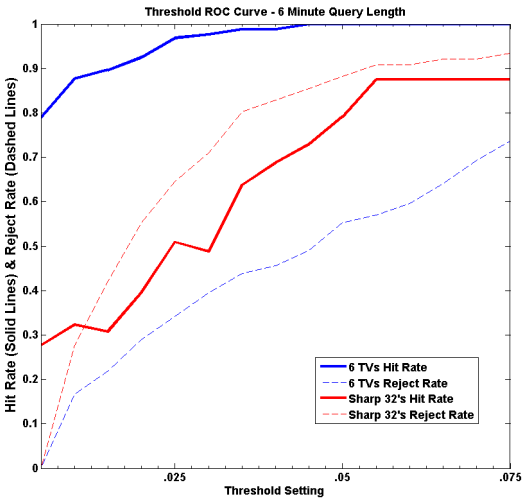
EMI signal fluctuations can be manifested as frequency shifts or variations in energy density. The following figure shows examples of these two types of EMI signals.



The Sharp-42 TV is shown on the left; changes in screen content cause a shift in the center frequency of the EMI. The white strip on left shows the signal obtained by tracking the EMI on left. The Panasonic-42-A TV is shown on the right; changes on screen manifest as change in the Energy density of the EMI. Tracking total energy over time yields the signal shown in white strip on the right.

B. REDUCING FALSE MATCHES USING A CONFIDENCE THRESHOLD

The following figure shows average hit rates (solid lines) and reject rates (dashed lines) as a function of threshold setting for confidence gap between the top match and the runner-up for 6 minute queries. The blue lines represent the average hit and reject rates for the 6 TVs (excluding the Sharp 32" A&B); and the red lines depict the average hit and reject rates for the Sharp 32" A&B. Note that the Sharp TVs produce significantly higher reject rates because they produce signals that are not good at discriminating between movies.



To clarify this figure, note that hit rate is *not* 1 minus the reject rate. Rather, the accept rate (not shown) is 1 minus the reject rate. The hit rate is the number of correct matches divided by the number of queries not rejected. The reject rate gets higher as we require a larger gap between the winner and the runner up. The accept rate (not shown) will simultaneously decrease. However, the hit rate (shown) does increase with larger threshold values. For our experiments we select a threshold of .05, which yields a hit rate.

C. HOME EMI

The following figure shows EMI recorded from Home 3. Note that multiple devices were in operation at any given time, and despite signal overlaps our EMI extraction algorithm was able to successfully track the TV noise signature.